LOCKSS

Lots (though Fewer) of (and More Hosted) Copies Keep Stuff Safe

Nicholas Taylor (@nullhandle)
Program Manager Emeritus, LOCKSS and
Web Archiving
Stanford Libraries

Internet Archive Web and Data Services Partner
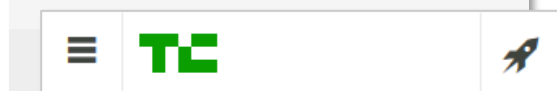Meeting
20 September 2019

What threatens the long-term persistence of digital information?

Amazon cloud crash wipes out customer data; Will users be compensated?

By **Molly McHugh** —
Posted on April 28, 2011 10:57 am

SECURITY

Ohio's Recent Spate of Cyberattacks Is Indicative of the National Trend

Angry Employee Deletes All of Company's Data

Published January 24, 2008

Fox News

Call it a tale of revenge gone wrong.

AddThis

Yahoo Quietly Pulls The Plug On Geocities

Posted Apr 23, 2009 by *Leena Rao* (*@leenarao*)

Sorry, new GeoCities accounts are no longer available.

Current GeoCities customers:

After careful consideration, we have decided to close GeoCities later this year. We'll share more details this summer. For now, please sign in or visit the help center for more information.

» Sign in to GeoCities

Save 50% on a web site with Yahoo! Web Hosting.

- Get a personalized web address (e.g., www.learn2yoga.com).
- Design a great-looking site with easy-to-use tools.
- Never worry about ads appearing on your site.
- Get unlimited disk space to store all your pictures and audio tracks.
- Welcome as many visitors as you like, thanks to unlimited data transfer. Details.
- Get answers even at 2 a.m. with 24-hour phone support.
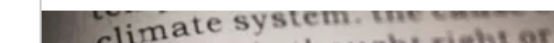
» Learn more | ▶ View demo

FUTURE PERFECT   EXPLAINERS   MORE ▾

"Climate change" and "global warming" are disappearing from government websites

The deletions follow a pattern of policy changes on climate change under the Trump administration.

By Umair Irfan | Updated Jan 11, 2018, 12:30pm EST

SHARE

climate system...

REUTERS

Davos   The Trump Effect   Politics

#SCIENCE NEWS
JULY 20, 2009 / 3:19 PM / 9 YEARS AGO

Moon landing tapes got erased, NASA admits

Maggie Fox, Health, Science Editor

WASHINGTON (Reuters) - The original recordings of the first humans landing on the moon 40 years ago were erased and re-used,

understand preservation threats

- long-term **data integrity** is hard

- need architecture informed by **actual leading threats** to data

- don't underestimate:
  - people making **mistakes**
  - **attacks** on information
  - organizational **failure**

"Fragile" by Garrett Coakley under CC BY-NC 2.0

LOCKSS

mitigate preservation threats

- lots of copies is **necessary but not sufficient**

- central points of failure can **undermine all copies at once**

- **multi-organizational** preservation storage provides:

  - **resilience** against organizational failure

  - **compartmentalization** against operator error

  - **decentralization** against attacks

  - **diversity** in technical infrastructure

LOCKSS

LOCKSS Networks

- a dozen **networks**
- hundreds of **institutions**
- **all types** of content
- preservation via **diversity** of:
  - technologies
 - institutions
  - networks
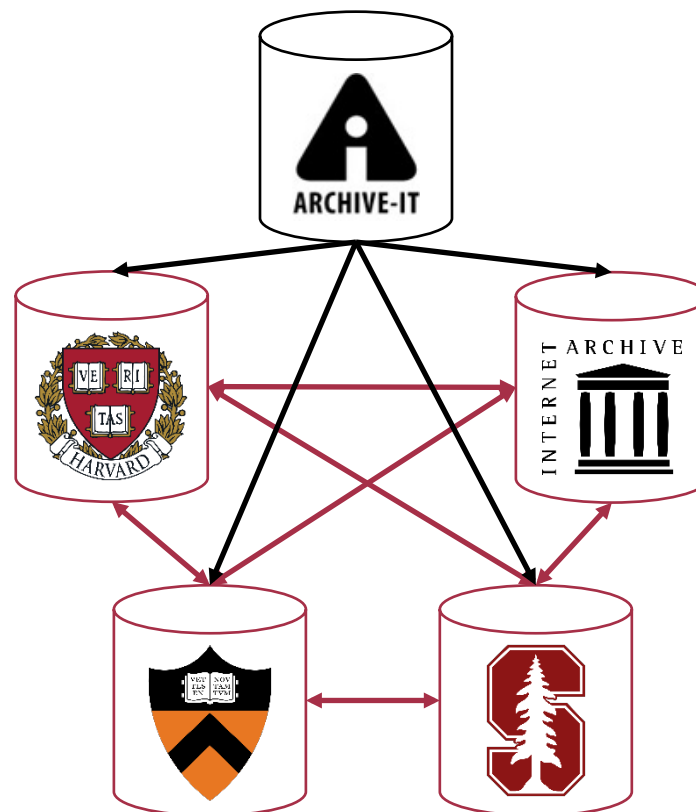

LOCKSS

network model evolution

- in original, Global LOCKSS Network, **all nodes stored copies**

- communities running private LOCKSS networks are **consolidating nodes**
  - **only a subset** run nodes
  - **whole community** is served, participates in governance, provides funding

- **under (co-)hosted service model**, Stanford + trusted partners may serve as anchor storage hosts

"Networks" by Alan Levine under CC 0
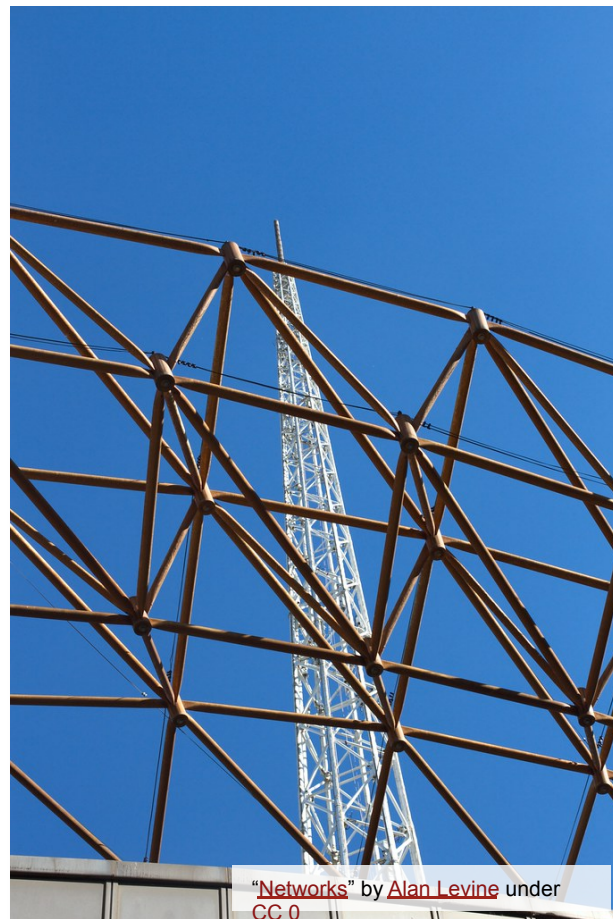
LOCKSS

IPLC Preservation Network pilot

- **5x nodes**: Harvard, IA (x2), Princeton, Stanford

- **all IPLC Archive-It collections** ingested + under active preservation

- ingest + audit/repair **take place automatically** via LOCKSS software

LOCKSS

Questions